International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(23s) 2026 pp. 604-621
https://doi.org/10.70917/ijcisim-2026-5627
Article

Design and Development of Hybrid Learning Model for the detection of Ovarian Cancers

V. Sathiyavani[1], Dr. Vinothkumar Arumugam[2], Dr. M. Anand[3], A Abilasha[4]

[1] Research Scholar, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India
[2],3 Professor, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India
[4]Assistant Professor, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India
sathiyavani25061983@gmail.com, dravinoth@gmail.com, anand.ece@drmgrdu.ac.in, vabilasha90@gmail.com

Abstract: Ovarian cancer remains as a significant contributor to female mortality across the globe, mainly because it has a difficulty in obtaining an early and precise diagnosis through complex histopathology images. Automated image-based diagnosis has therefore become essential to assist pathologists in making effective diagnoses of malignant tissues. Traditional diagnostic procedures and Deep Learning (DL) architectures often fail to capture fine-grained cellular textures as well as global contextual relationships needed for accurate classifications. These models are based on narrow receptive fields, have no focus on critical areas, and thus, the diagnostic accuracy and generalization across subtypes are limited. In order to overcome these shortcomings, the present research proposes a Hybrid Deep Learning Framework (ViT-SE + BiGRU-Luong FusionNet) to detect ovarian cancer accurately and effectively. The Vision Transformer (ViT) with a Squeeze-and-Excitation (SE) module is used to learn rich global contextual dependencies and focus important feature channels adaptively. Simultaneously, the Bidirectional Gated Recurrent Unit (BiGRU) with the Luong Attention mechanism accurately captures the sequential spatial dependencies and identifies critical pathological areas in tissue sequences. Finally, a Cross-Attention Fusion (CAF) module is utilised to combine complementary global and sequential representations to achieve robust classification. The proposed framework is trained and tested on the Kaggle Ovarian Cancer and Subtypes Histopathology Dataset and achieves an accuracy of 99.53%, precision of 99.41%, recall of 99.47%, specificity of 99.45% and F1-score of 99.44%. The experimental findings prove that the proposed hybrid framework substantially improves diagnostic accuracy and generalization without compromising high computational efficiency, which makes it very applicable in real-time clinical practice in medical image diagnosis.

Keywords: Ovarian Cancer, Deep Learning, Vision Transformer, Squeeze-and-Excitation, BiGRU, Luong Attention, Histopathology Images, Computational Efficiency

1. Introduction

Cancer persists as one of the major contributors to global mortality, and its early diagnosis is the key to a higher survival rate. Among many cancers that develop in women, ovarian cancer is the most dangerous one since it is very fatal and hard to detect at the early stages [1]. It is considered as one of the most aggressive gynecological cancers and a significant health issue among the population in the world [2]. In 2022, the World Cancer Research Fund stated that the rate of 324,603 new cases and 207,252 deaths make it the eighth-most common cancer in women [3]. Ovarian cancer is often called a silent killer because the initial symptoms are indefinitely vague, i.e. bloating or abdominal pain that can be easily ignored and contributes to the late diagnosis. Moreover, due to its biological diversity and the complexity of tumor structures, its timely and proper identification remains a persistent problem.

Ovarian cancer has a significant clinical, psychological, and socioeconomic effect, and patient survival depends greatly on the stage of cancer detection. The survival rate of early-stage diagnosis is approximately 90%, but detection at an advanced stage significantly reduces it to less than 30% [4]. Traditional methods of diagnosis, like ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), and blood biomarkers, e.g. CA-125, are still commonly used in the clinical practice. Nevertheless, these techniques are prone to shortcomings such as low

specificity, high cost, and expert judgment [5-7]. Though histopathological examination is the diagnostic gold standard, it is labor-intensive and prone to observer bias specifically in differentiating between benign and malignant forms. These shortcomings underscore the necessity of automated, objective and image-based diagnostic systems which can improve on accuracy and efficiency.

With the innovations of Machine Learning (ML) and DL architectures, computational diagnosis has transformed the field of cancer detection and classification [8,9]. DL models including Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and Transformers have exhibited impressive performance in the analysis of histopathological and radiological images of ovarian tumor. These models allow automatic features extraction, effective pattern recognition, and classification, which minimizes the human factor in clinical decisions [10]. The implementation of these intelligent models contributes to faster and evidence-based diagnostics, personal treatment planning, and reduced human error in cancer detection.

In spite of their advancement, the current models have a number of drawbacks in detecting ovarian cancer. CNN-based models are good at acquiring local textures but not global contextual dependencies, whereas transformer-based networks achieve good global representation but are computationally expensive [11]. Further, the majority of existing models do not prioritize key cancer-specific attributes, which may result in information overlap and inefficient predictions. The image quality, tissue structure variations, and imbalanced data also influence the model generalization [12]. Thus, it is essential to develop a hybrid DL framework to effectively combine local and global information enhancing accuracy, and computational efficiency to have practical medical diagnostic implementations.

1.1 Contribution of the Research

This study makes the following contributions for enhancing the application of intelligent DL methodologies in the detection of ovarian cancer:

1. A hybrid deep learning model is created by combining a ViT with SE module and a BiGRU with a Luong Attention mechanism to effectively extract both the global and sequential spatial details in histopathology images.
2. The CAF mechanism is constructed to incorporate the complementary feature representations from ViT-SE and BiGRU-Luong modules to achieve better feature interaction and enhanced performance on classification.
3. A comprehensive histopathological dataset of ovarian cancer and its subtypes is used, which has been preprocessed and normalized, as well as augmented to guarantee learning balance and robust model generalization.
4. The proposed framework is extensively experimentally and validated with the use of standard performance metrics (Accuracy, Precision, Recall, Specificity, and F1-score) to provide insights on the effectiveness, reliability, and computational efficacy of the suggested framework in comparison to the existing deep learning methods.

1.2 Paper Organization

The subsequent sections of this manuscript are outlined as follows: **Section 2** outlines the existing research background on the detection of ovarian cancer and also explains how DL algorithms has been applied to the analysis of histopathological images. **Section 3** discusses the data characteristics, preprocessing methods and the framework of the recommended hybrid system along with the Cross-Attention Fusion mechanism. **Section 4** details the experimental set up, performance testing, and a comparison against the conventional models. **Section 5** wraps up the study with the key findings, clinical implications of the model, and recommends the potential directions of future research.

2. Related Works

Macis et al. (2025) [13] developed a CNN-based tool for distinguishing endometriosis-associated ovarian cancer (EAOC) from non-EAOC and non-tumoral tissues utilising CE-CT scans. Four CNN models, including VGG19, Xception, ResNet50, and DenseNet121, were trained on axial slices with hybrid ensemble methods such as RF, GB, KNN, and SVM. The VGG19 + majority voting ensemble achieved an AUC of 0.85, while the VGG19 + SVM model reached an AUC of 0.97 for differentiating OC from non-tumoral cases. The study demonstrated the clinical potential of hybrid CNN-ML pipelines for precision surgery and early diagnosis. Performance evaluation included AUC, sensitivity, specificity, and MLcps scores. However, the framework remains limited by reliance on relatively small datasets and computational complexity for real-time adoption.

Garg et al. (2025) [14] provided a comprehensive review of ML approaches for recognizing gynecological cancers including breast, cervical, and ovarian cancers. The study highlighted the use of SVM, random forests, and CNNs for cancer type identification, risk classification, and treatment design. ML models showed strong capability in combining multi-modal data such as clinical, genomic, proteomic, and imaging features. Recent advances in explainable AI, federated learning, and multi-omics fusion were further examined to enhance applicability. The review emphasized improved precision oncology outcomes through ML adoption. Nonetheless, challenges remain with data inconsistency, interpretability issues, and integration into clinical workflows.

Adusumilli et al. (2025) [15] introduced a structured AI-enhanced diagnostic approach for identifying ovarian lesions from CT and MRI scans. The study employed both full-volume imaging pipelines and region-of-interest-focused processing strategies, utilizing advanced neural network architectures such as ResNet, DenseNet, a transformer-enhanced UNeST model, and an attention-driven Multiple Instance Learning framework. Experiments showed feasibility of the approach with transformers and MIL models achieving promising classification of ovarian masses. Class imbalance was handled with focal loss, oversampling, and dynamic class weighting. The PyTorch-based MONAI framework enabled reproducibility and optimization with Optuna. Yet, the framework is constrained by the preliminary dataset size and requires validation on larger multi-institutional data.

Tokareva et al. (2025) [16] designed a diagnostic pipeline for ovarian cancer that incorporates plasma lipidomics and metabolomics with ML algorithm. CNNs and XGBoost models with optimized feature selection achieved strong classification results across 229 plasma samples from patients and controls. CNN models attained an accuracy of 81% in classifying malignant and benign samples, whereas XGBoost combined with SVM-RFE yielded a 96% accuracy in distinguishing benign cases from healthy controls. For multiclass categorization, accuracies up to 78% were obtained using feature-selected CNN models. The research demonstrated the complementary role of ensemble and DL methods in biomarker discovery. However, clinical application is limited by the need for large-scale standardization and validation across diverse populations.

Ivanova et al. (2024) [17] introduced a label-free electrochemical biosensor for detecting lysophosphatidic acid (LPA), a biomarker for early ovarian cancer. Using a gold screen-printed electrode with a gelsolin–actin complex as a biorecognition element, the system quantified LPA concentrations via voltammetry. The biosensor demonstrated high sensitivity with an LOD of 0.9 μM and LOQ of 2.76 μM in goat serum samples. The study highlighted the potential of compact, cost-effective biosensors for point-of-care ovarian cancer diagnostics. Its findings support the feasibility of biosensor-based early-stage detection. Nevertheless, the limitation lies in the need for clinical validation in human patient samples and real-world settings.

Bucur et al. (2024) [18] reviewed the contribution of DL framework in ovarian cancers from diagnosis to treatment, focusing on biomarkers, imaging, and pathology integration. Their study highlighted how AI-driven systems can improve diagnostic precision, anticipate patient prognosis, and refine therapeutic planning. The review synthesized evidence from laboratory tests and imaging to show how AI supports both detection and prognosis. They concluded that AI-driven tools can significantly improve clinical workflows in ovarian cancer management. The results suggested a promising impact on early detection and patient survival. However, the limitation lies in the lack of standardized clinical integration and the need for further validation of AI-assisted systems.

Du et al. (2024) [19] proposed a multi-sequence feature fusion network using multi-parametric magnetic resonance imaging (mpMRI) for preoperative ovarian cancer subtype classification. Their deep learning method integrated multiple MRI sequences to distinguish high-grade serous carcinoma from clear cell carcinoma. The model achieved an accuracy of 91.62% and an AP of 95.13%, proving its efficiency in accurate subtype categorization. The fusion network demonstrated improved predictive power over traditional single-sequence models. This approach emphasized the importance of multi-parametric data in enhancing cancer diagnosis. Nevertheless, the model is limited by its dependence on large, high-quality mpMRI datasets for robust performance.

Ziyambe et al. (2023) [20] developed a CNN-based diagnostic model that analyzes histopathological images to detect ovarian cancer in pre- and post-menopausal patients. The CNN attained a diagnostic accuracy of 94%, correctly identifying 95.12% of cancer and 93.02% of healthy cells. The research demonstrated that AI-driven approaches can overcome human diagnostic challenges such as inter-observer variability and extended evaluation times. Their method provided a faster and more reliable diagnosis compared to conventional methods. The findings emphasized AI's role in improving cancer prediction efficiency. However, the limitation is the reliance on dataset diversity, as real-world generalization across institutions remains a challenge.

Jiang et al. (2023) [21] presented a detailed overview of DL-driven cancer diagnostic methods based on medical images, like X-ray, ultrasound, CT, MRI, PET, and histopathology. The study analyzed classical pretrained models alongside advanced networks like transformers, graph neural networks, and ensemble learning techniques. They focused on methods aimed at reducing overfitting, such as applying dropout, batch normalization, and augmenting the dataset. Applications of DL showed remarkable effectiveness in tasks like classification, object recognition, image segmentation, and registration. The results confirmed the growing value of AI in cancer diagnosis with improved accuracy and timeliness. Still, the drawback is the scarcity of high-quality annotated data and challenges in model generalization, especially for rare cancers.

Ghoniem et al. (2021) [22] suggested a multi-modal evolutionary DL framework incorporating genetic data and histopathological images for ovarian cancer diagnosis. Their model utilized predictive antlion-optimized LSTM for gene data and antlion-optimized CNN for histopathological images, followed by feature fusion. This hybrid architecture was benchmarked against nine other evolutionary fusion models, showing superior precision and accuracy across ovarian, breast, and lung cancer datasets. The results demonstrated the effectiveness of evolutionary optimization in enhancing multi-modal cancer diagnosis. Their approach outperformed existing methods in predictive accuracy and robustness. However, the limitation is the high computational cost and complexity of optimizing multi-modal architectures for real-world clinical use.

Table 1 presents a comparison of the existing related works, outlining their methodologies, key contributions, and limitations.

Table 1 Summary of the Related Works

S.No	Author & Year	Methodology	Advantage	Disadvantage
1	Macis et al. (2025)	CNN models (VGG19, Xception, ResNet50, DenseNet121) with hybrid ML ensembles (RF, GB, KNN, SVM); CE-CT axial slices; ensemble evaluation	VGG19 + SVM achieved AUC 0.97; potential for precision surgery and early diagnosis	Small dataset; high computational complexity limits real-time clinical adoption
2	Garg et al. (2025)	Review of ML models (SVM, RF, CNN); highlighted multi-modal fusion (clinical, genomic, proteomic, imaging);	Strong potential in precision oncology; improved treatment design & multi-omics integration	Data inconsistency; limited interpretability; clinical integration challenges

		discussed XAI & federated learning		
3	Adusumilli et al. (2025)	ResNet, DenseNet, transformer-based UNeST, Attention MIL; lesion-focused & whole-volume inputs; MONAI framework with Optuna optimization	Promising performance with transformers & MIL; reproducible workflows; robust imbalance handling	Preliminary dataset size; requires validation on larger, multi-institutional datasets
4	Tokareva et al. (2025)	CNN & XGBoost with SVM-RFE feature selection; 229 plasma samples; ensemble methods for biomarker discovery	CNN achieved 81% accuracy; XGBoost+SVM-RFE reached 96%; strong biomarker identification	Needs large-scale validation and standardization across diverse populations
5	Ivanova et al. (2024)	Gold screen-printed electrode with gelsolin–actin biorecognition; voltammetry; tested in goat serum samples	High sensitivity (LOD = 0.9 μM); compact & cost-effective; point-of-care diagnostic potential	Requires clinical validation in human samples and real-world testing
6	Bucur et al. (2024)	Literature review on AI applications across biomarkers, imaging, pathology, and treatment optimization	Showed AI's role in detection, prognosis, and workflow optimization; potential for survival benefit	Lack of standardized clinical integration; validation of AI tools needed
7	Du et al. (2024)	Deep learning fusion of multi-parametric MRI sequences; distinguished serous vs. clear cell carcinoma	High accuracy (AUC 91.62%, AP 95.13%); better than single-sequence models	Dependent on large, high-quality mpMRI datasets
8	Ziyambe et al. (2023)	CNN model on histopathological images; cancer vs. healthy classification	High accuracy (94%); reduced inter-observer variability; faster, reliable diagnosis	Limited dataset diversity; generalization across institutions remains a challenge
9	Jiang et al. (2023)	Review of pretrained DL models (CNNs, transformers, GNNs, ensembles); applied to X-ray, CT, MRI, PET, histopathology	Improved accuracy & timeliness across multiple cancer imaging tasks	Scarcity of annotated datasets; poor generalization for rare cancers
10	Ghoniem et al. (2021)	Antlion-optimized LSTM (genomic) + CNN (histopathology); feature fusion; compared against 9 fusion models	Superior accuracy & robustness; effective evolutionary optimization in multi-modal diagnosis	High computational cost; complex multi-modal optimization hinders clinical deployment

3. Proposed Framework

Figure 1 depicts the overall framework of the recommended hybrid model for ovarian cancer detection. The framework integrates global spatial features from the ViT-SE branch and sequential dependencies from the BiGRU-Luong branch, which are fused through the Cross-Attention Fusion (CAF) mechanism. This fusion enables comprehensive feature representation for accurate cancer subtype classification.

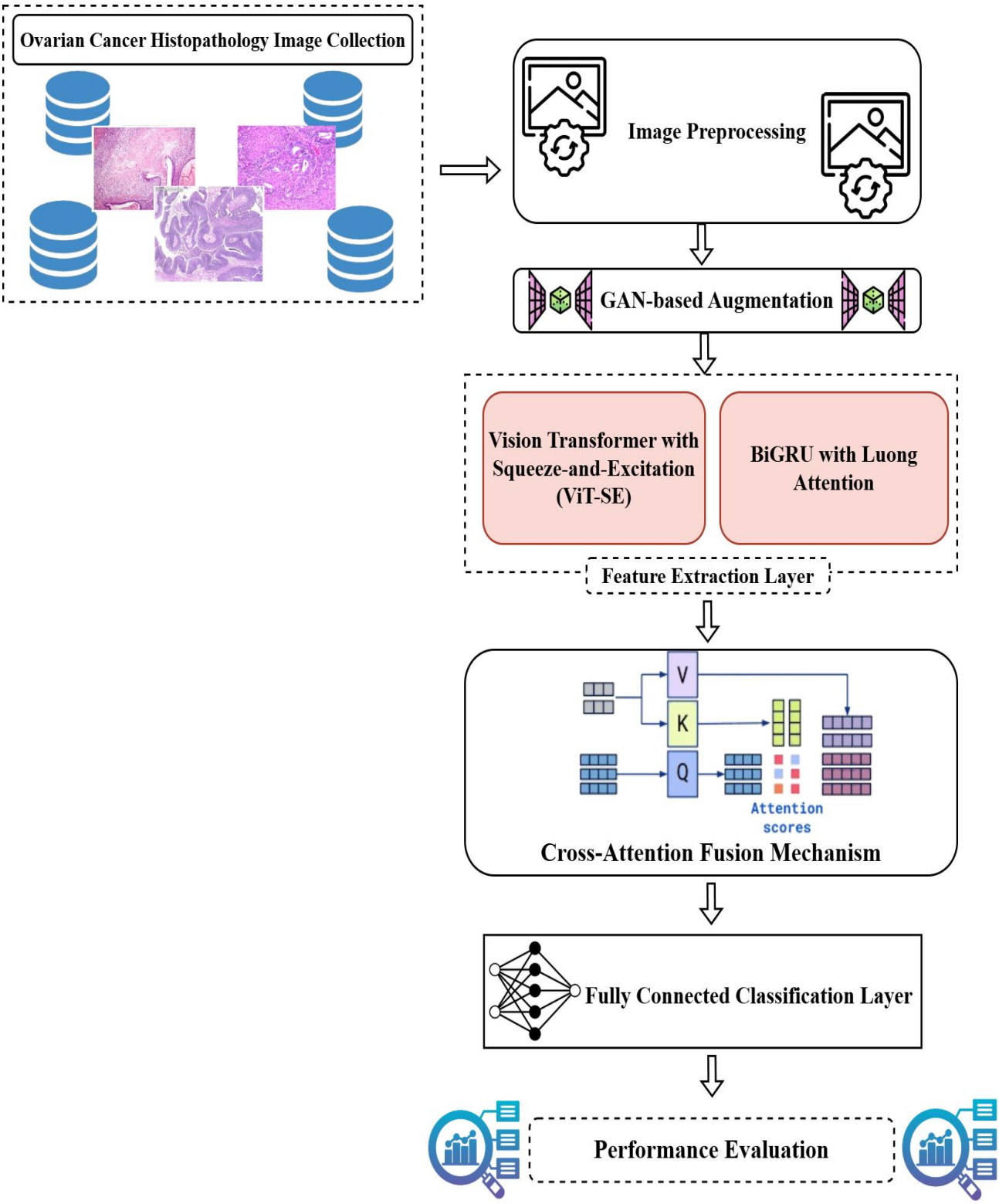
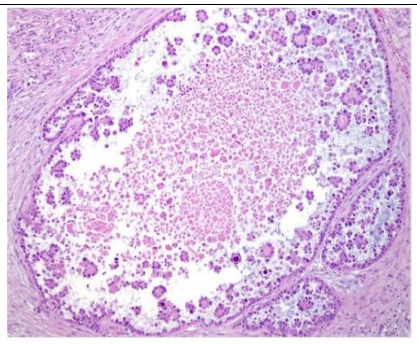
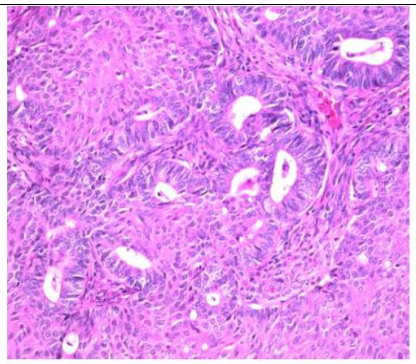
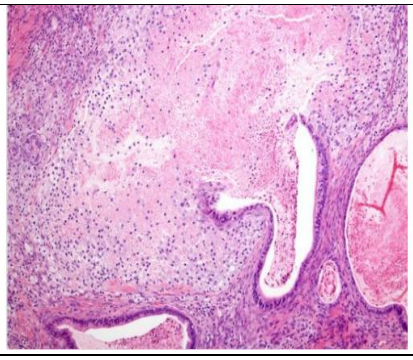


Figure 1 Comprehensive Layout of the Recommended Approach

3.1 Materials and Methods

This research utilizes the **Ovarian Cancer and Subtypes Histopathology Dataset**, which is publicly available on Kaggle [23]. The dataset offers a comprehensive collection of histopathology images of both cancerous and non-cancer tissue samples of the ovaries. It comprises four major subtypes of ovarian cancer such as Clear Cell Carcinoma, Endometrioid Carcinoma, Mucinous Carcinoma and Serous Carcinoma as well as a distinct class of Non-Cancerous tissue images. All images are obtained by using microscopic histopathology slides with Hematoxylin and Eosin(H&E) staining, which involves capturing the cellular morphology and tissue organization patterns that are important in the study of pathology. The dataset contains 498 images, which are well organized in separate directories per subtype group, assuring clear labeling for classification tasks. Images are stored in JPEG format with high spatial resolutions and stable color values, which makes it possible to extract texture and structural characteristics with high precision and analyze them with deep learning techniques. Figure 2 shows a sample histopathological image of each ovarian cancer subtypes and the non- cancerous image.

Types	**Sample Image**
Clear Cell Carcinoma	
Endometrioid Carcinoma	
Mucinous Carcinoma	

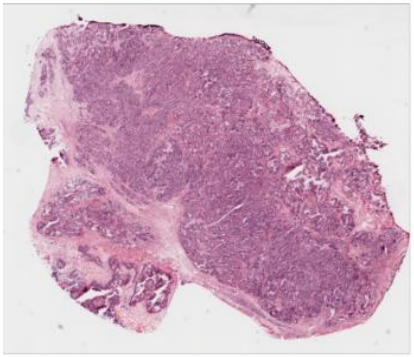
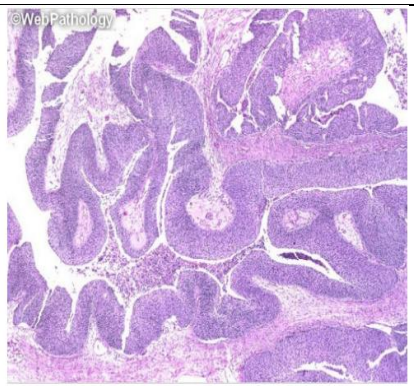
Serous Carcinoma	
Non-Cancerous	

Figure 2 Sample images of each class from the dataset.

3.2 Data Preprocessing

Prior to training, a number of preprocessing and augmentation methods were used to enhance the quality of images, normalize data distribution, and alleviate class imbalance in the dataset. First, all histopathological images were scaled to 224 × 224 pixels in order to preserve similar dimensions, and Gaussian filtering was used to decrease the amount of noise and even out the intensity of color between samples. To further improve the diversity of the data and deal with the class imbalance, Generative Adversarial Networks (GANs), were employed to produce realistic synthetic histopathology images that retained the cellular and structural features of every images in the dataset. Furthermore, Synthetic Minority Oversampling Technique (SMOTE) was used at feature-level to equalize the minority classes, so that synthetic feature vectors were generated, thus making all the categories equally represented. Table 2 presents the augmented dataset distribution, which included 10,000 histopathological images in five categories. These data were split into 70% training and 30% testing to provide equal assessment of the proposed framework.

Table 2 Dataset Distribution after Augmentation and Preprocessing

Class (Subtype)	**Training Set (70%)**	**Testing Set (30%)**
Clear Cell	1,400	600
Endometrioid	1,400	600
Mucinous	1,400	600
Serous	1,400	600
Non-Cancerous	1,400	600

3.3 Model Design

This section discusses about the suggested Vision Transformer (ViT), Squeeze-and-Excitation (SE) module, Bidirectional Gated Recurrent Unit (BiGRU), and Luong Attention mechanism. A detailed description of each component is provided below.

3.3.1 ViT

The ViT is a DL algorithm that applies the transformer-based self-attention strategy, initially built for NLP, to image recognition tasks. Unlike convolutional networks, ViT divides an image into fixed-size patches and treats them as tokens in a sequence, thereby capturing long-range dependencies between different image regions. This structure enables ViT to model global contextual relationships, which are crucial for distinguishing subtle histopathological patterns in ovarian cancer subtypes. Figure 3 illustrates the general structure of the ViT model.

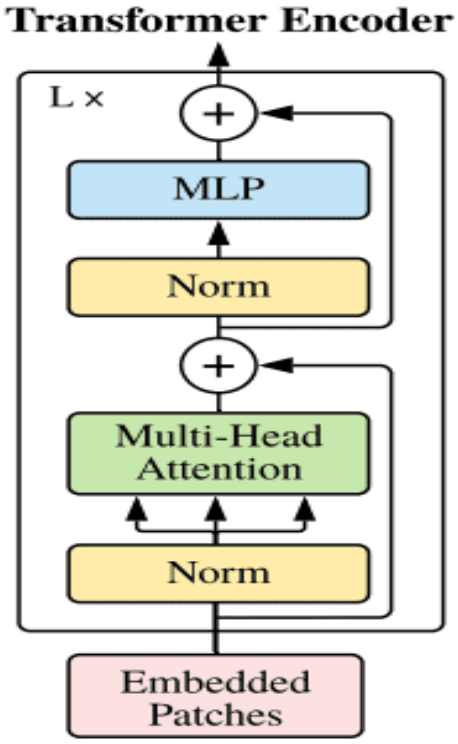


Figure 3 Structure of the ViT model

Each image $I \in RH \times W \times C$ is divided into non−overlapping image patches of size $P \times P$, here H and W denote the height and width of the image, and C represents the count of channels. The total count of patches N is given by:

$$N = \frac{H \cdot W}{P^2} \qquad (1)$$

Each image patch is transformed into a one-dimensional vector and mapped into a D-dimensional feature space via the embedding matrix E. To preserve positional information, positional embeddings E_{pos} are appended to every projected patch. A learnable class token z_{cls} is concatenated with the embedded patch sequence, resulting in the initial input to the Transformer encoder:

$$Z^0 = [z_{cls}; x_1E; x_2E; \ldots; x_NE] + E_{pos} \qquad (2)$$

The encoder contains several identical layers, consisting of a Multi-Head Self-Attention (MHSA) block and a Multi-Layer Perceptron (MLP) block. For each layer l, the input sequence z_{l-1} is first normalized using Layer Normalization (LN) and then processed by the attention block. Residual connections are applied after both sublayers to enhance gradient flow and improve stability:

$$z_l' = MHSA(LN(z_{l-1})) + z_{l-1} \quad (3)$$

$$z_l = MLP(LN(z_l')) + z_l' \quad (4)$$

The MHSA mechanism computes attention weights between the query Q, key K, and value V matrices as:

$$Attention(Q, K, V) = Softmax(\frac{QK^T}{\sqrt{d_k}}) \cdot V \quad (5)$$

Where, d_k represents the size of the key vectors, and the Softmax function normalizes the attention scores so that their total equals one.

This enables the network to focus on critical global regions of the image by assigning adaptive attention scores to significant tissue patterns. The ViT effectively captures long-range spatial dependencies and complex contextual interactions across image patches, making it highly suitable for differentiating histopathological patterns in ovarian cancer images.

3.3.2 Squeeze-and-Excitation (SE) Module

Although ViT efficiently models spatial dependencies, it lacks the ability to adaptively emphasize the most informative feature channels. To overcome this limitation, a SE mechanism is integrated with ViT to enhance the representational power by recalibrating channel-wise feature responses. The SE mechanism is particularly effective in highlighting discriminative patterns by learning which channels contribute most to the classification task. By doing so, it suppresses irrelevant or redundant channels while amplifying those that hold meaningful pathological information. The SE module operates through two core phases: squeeze and excitation.

In the squeeze operation, global average pooling aggregates spatial information from all channel into a channel descriptor as represented by the below equation,

$$z_c = \frac{1}{H \cdot W} \cdot \sum_{i=1}^{H} \quad \sum_{j=1}^{W} \quad X_c(i,j) \quad (6)$$

Here $X_c(i,j)$ represents the value at location (i,j) of the c-th feature map, H and W are the spatial height and width of the feature map, z_c symbolizes the channel descriptor summarizing the contribution of the c-th channel

In the excitation phase, two fully connected layers with a ReLU and sigmoid activation generate adaptive channel weights that model inter-channel dependencies:

$$s = \sigma(W_2 \cdot \delta(W_1 \cdot z)) \quad (7)$$

Here, δ represents the ReLU activation function, σ is the sigmoid function, and W_1 and W_2 signifies learnable weight matrices. The recalibrated feature map is then obtained through element-wise multiplication:

$$\widetilde{X_c} = s_c \times X_c \quad (8)$$

Here, s_c is the learned importance weight for channel c, and it scales the original feature map X_c to emphasize or suppress that channel. By integrating the SE mechanism within the ViT branch, the model adaptively emphasizes the most discriminative feature channels while preserving global contextual understanding. This combination enhances both spatial and channel-level feature representation, enabling more accurate and robust classification of ovarian cancer subtypes.

3.3.3 BiGRU

The BiGRU algorithm enhances a traditional GRU by analyzing the input data in both forward and reverse directions. This bidirectional flow enables the network to access both past and future contextual cues, allowing it to capture richer structural and semantic dependencies within the feature sequence. Thus, BiGRU is highly effective in identifying fine-grained textural and morphological patterns in tissue regions, improving the reliability and discriminative capacity of the learned representations. Figure 4 illustrates the architecture of BiGRU network model.

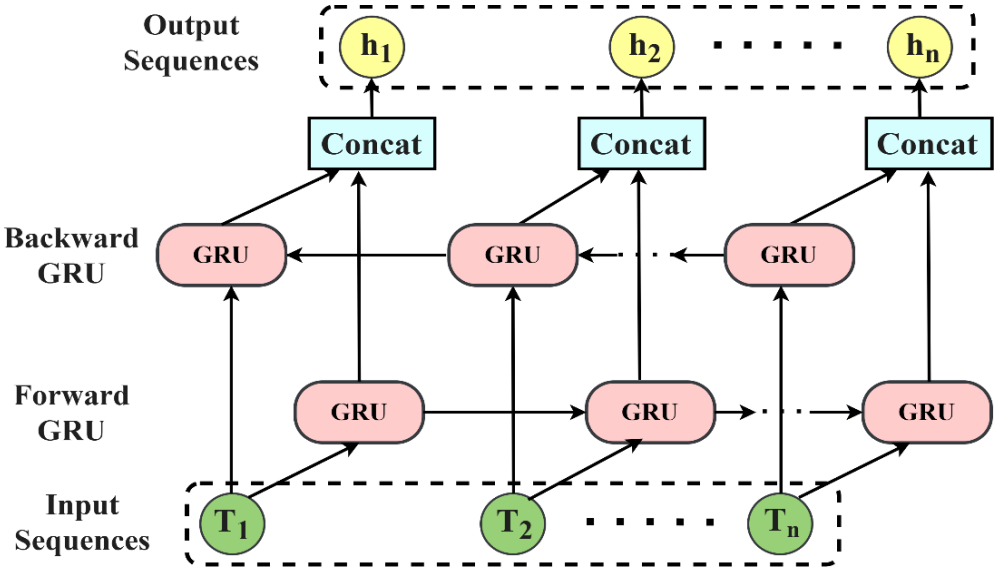


Figure 4 BiGRU Network architecture

At each time step, these gates manage how much historical information is carried forward and how much of the current input is processed. The update gate is calculated as:

$$Z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (9)$$

It manages how strongly past hidden features influence the current state h_{t-1} is retained. The reset gate regulates how much previous information is ignored:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (10)$$

A candidate hidden state $\hat{h_t}$ is generated using the reset gate's influence:

$$\hat{h_t} = tanh\ tanh\ (W_h \cdot [r_t \cdot h_{t-1}, x_t] + b_h) \quad (11)$$

Finally, the current hidden state h_t is updated as:

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \hat{h_t} \quad (12)$$

Where, x_t represents the input vector at time step t, h_{t-1} represents the hidden state from the previous time step, W_z, W_z, W_h refer to the weight matrices, and b_z, b_r, b_h correspond to their respective bias vectors. The sigmoid function σ and hyperbolic tangent methods $tanh$ are used for non-linear transformations, allowing the network to capture and model complex dependencies within the data.

3.3.4 Luong Attention

The Luong Attention mechanism is integrated with the BiGRU branch to refine the feature representation by dynamically weighting the contribution of each hidden state. Although BiGRU effectively captures bidirectional

dependencies within histopathological sequences, not all time steps contribute equally to the final prediction. Luong Attention enables the model to focus more on the most relevant temporal positions, improving interpretability and classification accuracy.

Let h_t denote the current decoder hidden state and h_s represent the encoder hidden states. The attention score between them is computed as:

$$score(h_t, h_s) = h_t^T W_a h_s \quad (13)$$

Here W_a is a learnable weight matrix that aligns the relevance between h_t and h_s. The alignment weights are obtained using a softmax function:

$$\alpha_t = \frac{exp(score(h_t h_s))}{\sum s' exp(score(h_t h_s'))} \quad (14)$$

These weights quantify how much attention should be given to every hidden state h_s when generating the present output. The context vector c_t is then derived as a weighted sum of encoder hidden states:

$$c_t = \sum_s \quad \alpha_t h_s \quad (15)$$

Finally, the attended context c_t is concatenated with the current hidden state h_t to form an enhanced representation:

$$\widetilde{h}_t = tanh(W_c[c_t; h_t]) \quad (16)$$

Here W_c denotes a trainable parameter matrix. This enriched hidden state $\widetilde{h}_t$captures both the sequential dependencies and the relative significance of each temporal element. By incorporating Luong Attention, the framework adaptively emphasizes critical cellular patterns across timestep, leading to more robust and context-aware recognition of ovarian cancer subtypes.

3.4 CAF Mechanism

This module acts as the central integration layer that unifies the high−level spatial features learned by the ViT−SE branch with the temporal dependencies captured by the BiGRU−Luong module. While the ViT−SE encoder extracts global texture and structural representations from histopathological images, the BiGRU−Luong network models contextual and sequential correlations. The CAF mechanism enables effective interaction between these heterogeneous representations through a cross−attention operation. Let the feature map from the ViT−SE branch be denoted as Fvit∈RN×D and the contextual embedding from BiGRU−Luong as Fgru∈RM×D. In the CAF block, the ViT features act as queries, while the BiGRU features serve as keys and values, allowing the model to align sequential dependencies with spatial semantics.

The attention alignment matrix is computed as:

$$A = softmax(\frac{QK^T}{\sqrt{d}}) \quad (17)$$

where Q=WQFvit, K= WkFgru, and d is the feature dimension used for scaling. This operation determines the degree of relevance between ViT and BiGRU feature domains. The cross-attended feature representation is then derived as:

$$F_{caf} = AV, where\ V = W_V F_{gru} \quad (18)$$

Here, F_{caf} combines the spatial and sequential dependencies, enabling the model to focus on complementary aspects of both feature types. To ensure balanced integration, a gating coefficient α is applied to regulate the contribution of each branch:

$$F_{final} = \alpha F_{caf} + (1 - \alpha) F_{vit} \quad (19)$$

At last, the fused feature F_{final} is passed through normalization and a fully connected layer for classification. The CAF mechanism strengthens inter-domain interactions, enhances gradient stability, and increases feature diversity, resulting in a more discriminative and robust representation for ovarian cancer subtype classification.

4. Implementation Details

The suggested framework in this study was built by Keras with tensorflow as backend. During the training phase**,** early stopping**,** along with dropout and L2 regularization**,** was applied to prevent overfitting and achieve stable convergence. Five-fold cross-validation was employed for model validation and testing to ensure the reliability and generalization capability of the proposed approach. The proposed model was also tested in a cloud computing environment (Google Colab Pro+), utilizing NVIDIA A100 Tensor Core GPU, Intel Xeon 2.8 GHz CPU, and 25 GB of system memory, enabling scalable distributed training and rapid experimental iterations.

4.1 Evaluation Measures

The efficiency of the suggested ViT-SE + BiGRU-Luong FusionNet framework was assessed utilising standard classification measures like accuracy, precision, recall, specificity, and F1-score. These metrics were employed to comprehensively assess the model's effectiveness in classifying the five target categories with high precision and reliability.

4.2 Result Outcomes

The proposed ViT-SE + BiGRU-Luong FusionNet model was evaluated using 3000 samples from the 30% test set. The confusion matrix in Figure 5 demonstrates the algorithm's exceptional predictive performance, achieving an overall accuracy of 99.53%, correctly categorizing 2986 samples and misclassifying only 14 samples across the five performance categories (Class 0–Class 4). The training and validation accuracy curves, illustrated in Figure 7, show a consistent upward trend and smooth convergence over 250 epochs, indicating efficient learning and strong generalization. Furthermore, the ROC–AUC curves shown in Figure 6 confirm the model's superior discriminative capability across all five classes. All experiments were conducted using an 70/30 train–test split, with a fixed random seed of 42 to ensure reproducibility and robust performance validation.

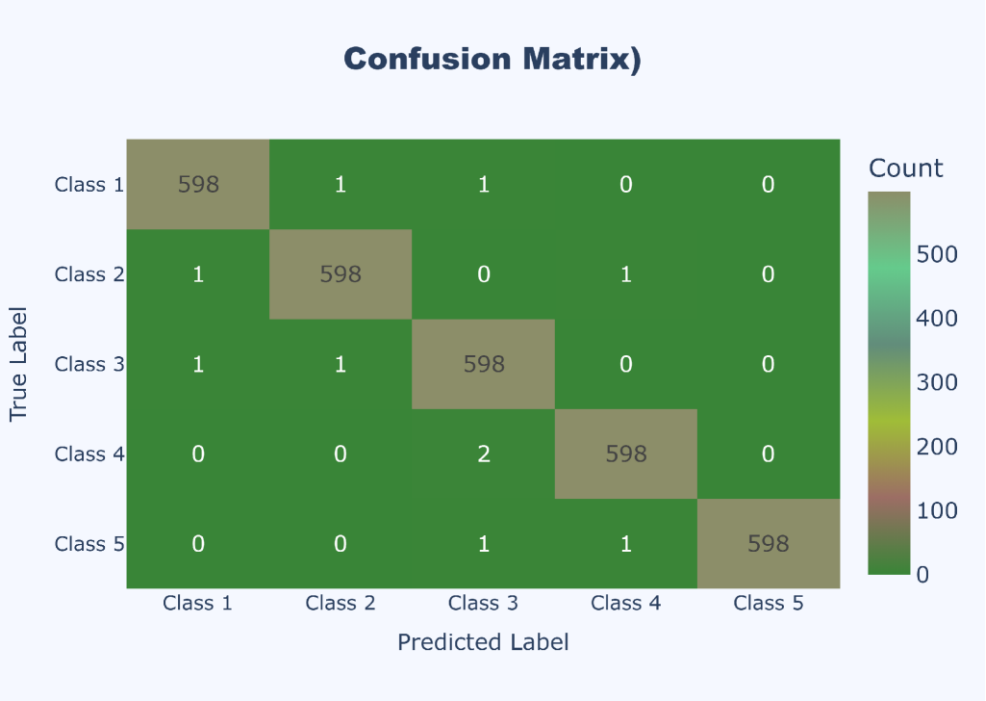


Figure 5 Confusion matrix shows the classification performance of the Suggested ViT-SE + BiGRU-Luong FusionNet across five classes

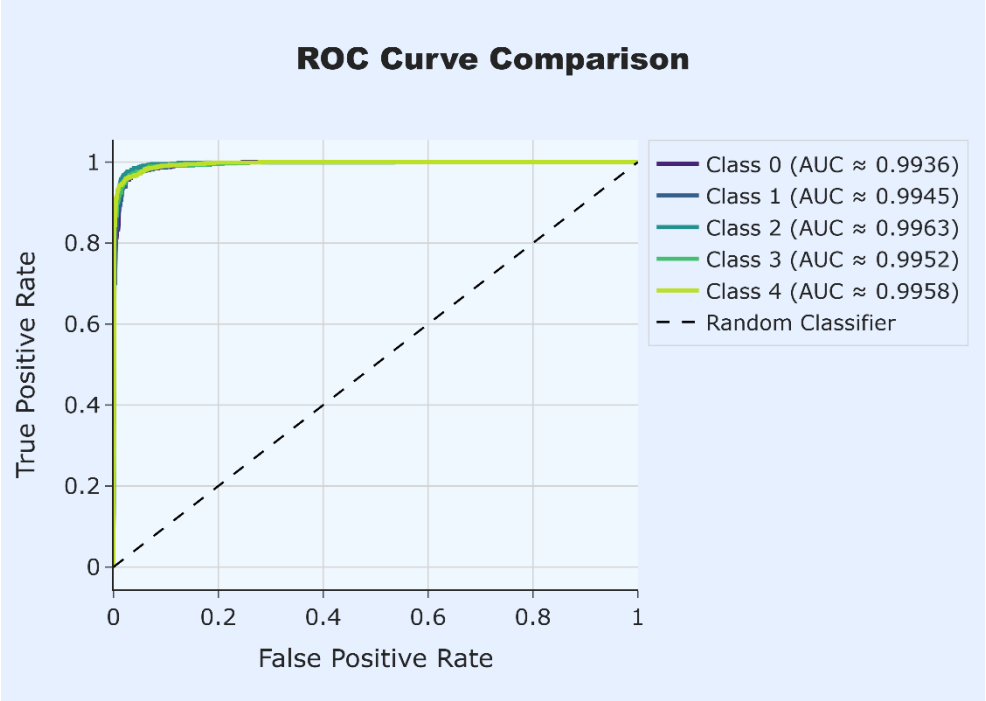


Figure 6 ROC–AUC curves of the proposed system

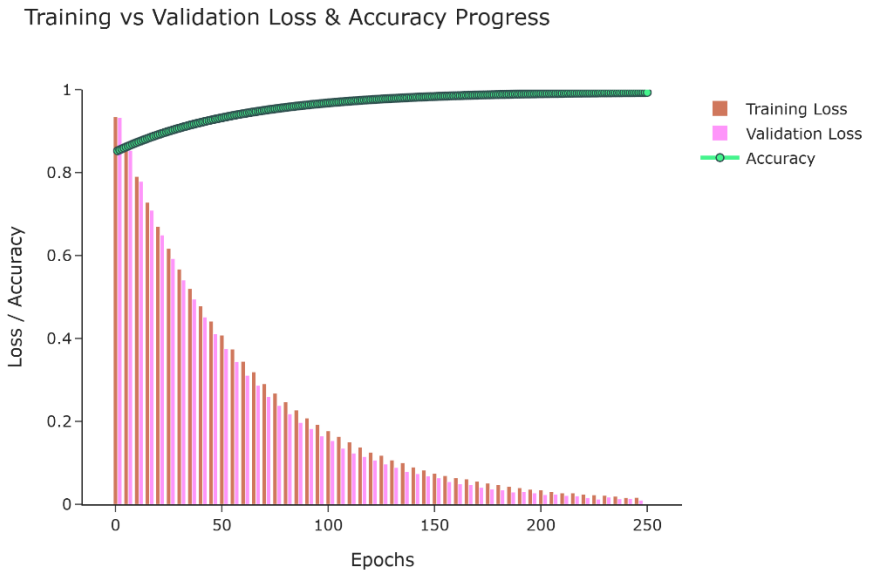


Figure 7 Training and validation loss curves of the proposed approach over 250 epochs, showing smooth convergence

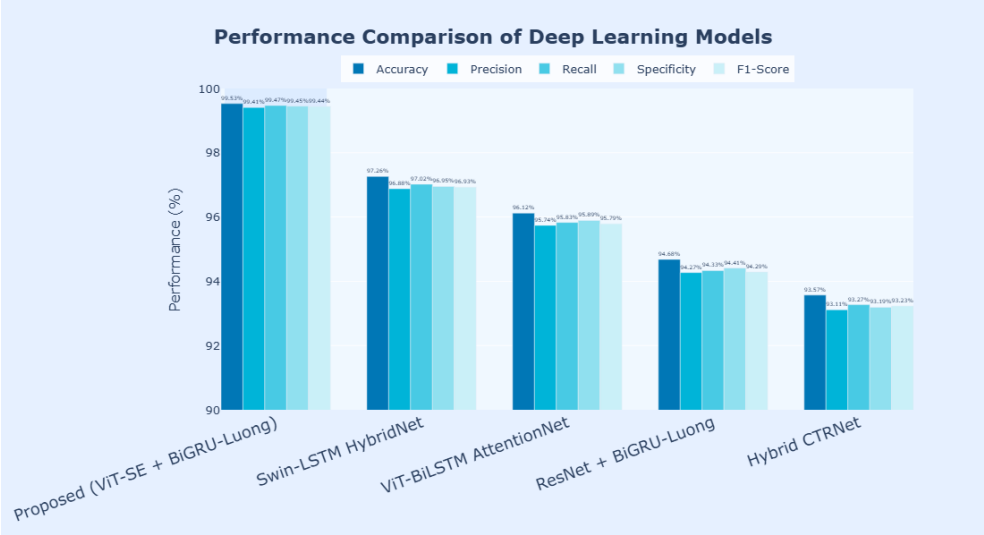


Figure 8 Comparative analysis showing the superior accuracy of the proposed ViT-SE + BiGRU-Luong FusionNet over existing models.

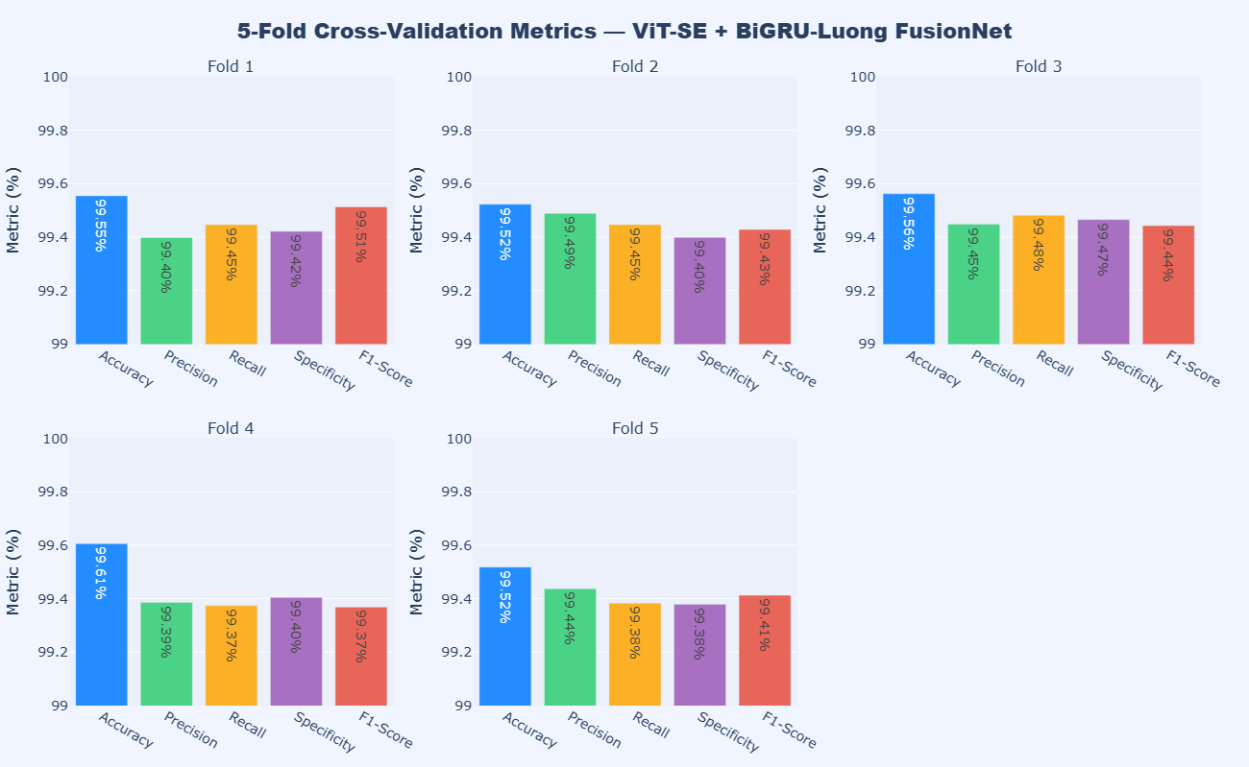


Figure 9 Five-fold cross-validation results demonstrating the consistent performance and robustness of the suggested model.

The effectiveness of the recommended ViT-SE + BiGRU-Luong FusionNet was extensively evaluated on the same dataset under identical conditions. As shown in Figure 8, the recommended algorithm achieved the highest performance, obtaining an accuracy of 99.53%, precision of 0.994, recall of 0.995, specificity of 0.994 and an F1-score of 0.994, demonstrating exceptional classification capability and stable convergence. To further validate its robustness and consistency, a five-fold cross-validation was performed, as illustrated in Figure 9, showing minimal performance variation across folds. The low standard deviation confirms the model's high stability and strong generalization ability, proving its effectiveness in learning complex spatial–temporal relationships with remarkable accuracy and reliability. To analyze the contribution of each major component, an ablation study was conducted using five configurations derived from the proposed ViT-SE + BiGRU-Luong FusionNet. Each variant was trained under identical settings to ensure fair comparison. The findings in Table 3 highlight the individual effect of Vision Transformer (ViT), Squeeze-and-Excitation (SE), Bidirectional GRU (BiGRU), and Luong Attention modules.

Table 3 Ablation study illustrating the performance impact of each component in the suggested ViT-SE + BiGRU-Luong FusionNet.

Model Variant	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ViT only	95.72	95.34	95.45	95.38
ViT + SE	96.45	96.18	96.26	96.21
BiGRU only	93.36	92.97	93.12	93.04
ViT-SE + BiGRU (No Luong)	97.86	97.57	97.63	97.59
ViT-SE + BiGRU-Luong (Proposed)	**99.53**	**99.41**	**99.47**	**99.44**

The results show that SE integration enhances feature discrimination within ViT, BiGRU contributes temporal dependency modeling, and Luong Attention further aligns contextual relations for precise decision-making. The recommended algorithm achieves the highest accuracy (99.53%) and F1-score (99.44%), confirming the efficacy of combining spatial–temporal attention and feature fusion.

5. Conclusion

This research article presents a robust and intelligent hybrid DL architecture for ovarian cancer recognition based on histopathological images. By incorporating the ViT with SE module and a BiGRU network with Luong Attention, the model can overcome the drawbacks of traditional diagnostic and deep learning techniques. The ViT-SE branch can effectively capture global contextual dependencies and prioritize important channel features, whereas the BiGRU-Luong component can capture sequential spatial dependencies and underscore critical pathological locations. This complementary design allows the model to discriminate malignant tissue patterns which attained an accuracy of 99.53%, precision of 99.41%, recall of 99.47%, specificity of 99.45% and F1-score of 99.44%. The suggested framework provides high levels of generalization, computing performance, and flexibility to various ovarian cancer subtypes. Its performance presents a promising step towards real-time, automated histopathological diagnosis, which could be used in clinical decision-making and improving the reliability of pathologist's diagnosis. Future enhancement may focus to the expansion of the model to multi-class cancer grade, the incorporation of multimodal medical data, and the implementation of the model in real-world diagnosis procedures to promote precision oncology.

6. Acknowledgement:

The author, Mrs. V. Sathiyavani, Associate Professor, Department of ECE, Kuppam Engineering College, Kuppam , sincerely acknowledges Kuppam Engineering College for providing the necessary research facilities and support through its research centre, which enabled and facilitated the successful completion of this research work. The author is grateful for the institution's continued encouragement and support.

REFERENCES

1. Matsas, A., Stefanoudakis, D., Troupis, T., Kontzoglou, K., Eleftheriades, M., Christopoulos, P., Panoskaltsis, T., Stamoula, E., & Iliopoulos, D. C. (2023). Tumor Markers and Their Diagnostic Significance in Ovarian Cancer. *Life*, *13*(8), 1689. https://doi.org/10.3390/life13081689
2. Santoro, A., Angelico, G., Travaglino, A., Inzani, F., Spadola, S., Pettinato, A., Mazzucchelli, M., Bragantini, E., Maccio, L., & Zannoni, G. F. (2025). *The multiple facets of ovarian high-grade serous carcinoma: A review on morphological, immunohistochemical and molecular features*. *Critical Reviews in Oncology/Hematology, 208*, 104603. https://doi.org/10.1016/j.critrevonc.2024.104603
3. Kumar, N., Mangla, M. Nanotechnology and nanobots unleashed: pioneering a new era in gynecological cancer management – a comprehensive review. *Cancer Chemother Pharmacol* 95, 18 (2025). https://doi.org/10.1007/s00280-024-04747-4
4. Holcakova, J., Bartosik, M., Anton, M., Minar, L., Hausnerova, J., Bednarikova, M., Weinberger, V., & Hrstka, R. (2021). New Trends in the Detection of Gynecological Precancerous Lesions and Early-Stage Cancers. *Cancers*, *13*(24), 6339. https://doi.org/10.3390/cancers13246339
5. Malik, M. M. U. D., Alqahtani, M. M., Hadadi, I., Kanbayti, I., Alawaji, Z., & Aloufi, B. A. (2024). Molecular Imaging Biomarkers for Early Cancer Detection: A Systematic Review of Emerging Technologies and Clinical Applications. *Diagnostics*, *14*(21), 2459. https://doi.org/10.3390/diagnostics14212459
6. Das, S., Dey, M. K., Devireddy, R., & Gartia, M. R. (2024). Biomarkers in Cancer Detection, Diagnosis, and Prognosis. *Sensors*, *24*(1), 37. https://doi.org/10.3390/s24010037
7. Oyowvi, M.O., Babawale, K.H., Atere, A.D. *et al.* Emerging nanotechnologies and their role in early ovarian cancer detection, diagnosis and interventions. *J Ovarian Res* **18**, 96 (2025). https://doi.org/10.1186/s13048-025-01678-3
8. Ahamad, M. M., Aktar, S., Uddin, M. J., Rahman, T., Alyami, S. A., Al-Ashhab, S., Akhdar, H. F., Azad, A., & Moni, M. A. (2022). Early-Stage Detection of Ovarian Cancer Based on Clinical Data Using Machine Learning Approaches. *Journal of Personalized Medicine*, *12*(8), 1211. https://doi.org/10.3390/jpm12081211
9. Juwono, F. H., Wong, W. K., Pek, H. T., Sivakumar, S., & Acula, D. D. (2022). *Ovarian cancer detection using optimized machine learning models with adaptive differential evolution*. *Biomedical Signal Processing and Control, 77*, 103785. https://doi.org/10.1016/j.bspc.2022.103785

10. Kongara, S.R., Prakasha, S., Brindha, A. et al. Performance evaluation of optimized convolutional neural network mechanism in the detection and classification of ovarian cancer. Multimed Tools Appl 83, 71311–71334 (2024). https://doi.org/10.1007/s11042-024-18115-0
11. Li, G., Zhang, Y., Li, K., Liu, X., Lu, Y., Zhang, Z., Liu, Z., Wu, Y., Liu, F., Huang, H., Yu, M., Yang, Z., Zheng, X., Guo, C., Gao, Y., Wang, T., Fok, M., Lau, J. Y.-N., Shi, K., Gu, X., Guo, L., Luo, H., Zeng, F., & Zhang, K. (2024). Transformer-based AI technology improves early ovarian cancer diagnosis using cfDNA methylation markers. Cell Reports Medicine, 5(8), 101666. https://doi.org/10.1016/j.xcrm.2024.101666
12. M. Radhakrishnan, N. Sampathila, H. Muralikrishna and K. S. Swathi, "Advancing Ovarian Cancer Diagnosis Through Deep Learning and eXplainable AI: A Multiclassification Approach," in *IEEE Access*, vol. 12, pp. 116968-116986, 2024, doi: 10.1109/ACCESS.2024.3448219.
13. Macis, C., Santoro, M., Zybin, V., Di Costanzo, S., Coada, C. A., Dondi, G., De Iaco, P., Perrone, A. M., & Strigari, L. (2025). A Convolutional Neural Network Tool for Early Diagnosis and Precision Surgery in Endometriosis-Associated Ovarian Cancer. *Applied Sciences*, *15*(6), 3070. https://doi.org/10.3390/app15063070.
14. Garg, P., Krishna, M., Kulkarni, P., Horne, D., Salgia, R., & Singhal, S. S. (2025). Machine Learning Models for Predicting Gynecological Cancers: Advances, Challenges, and Future Directions. *Cancers*, *17*(17), 2799. https://doi.org/10.3390/cancers17172799.
15. Adusumilli, P., Ravikumar, N., Hall, G., & Scarsbrook, A. F. (2025). A Methodological Framework for AI-Assisted Diagnosis of Ovarian Masses Using CT and MR Imaging. *Journal of Personalized Medicine*, *15*(2), 76. https://doi.org/10.3390/jpm15020076.
16. Tokareva, A., Iurova, M., Starodubtseva, N., Chagovets, V., Novoselova, A., Kukaev, E., Frankevich, V., & Sukhikh, G. (2025). Machine Learning Framework for Ovarian Cancer Diagnostics Using Plasma Lipidomics and Metabolomics. *International Journal of Molecular Sciences*, *26*(14), 6630. https://doi.org/10.3390/ijms26146630.
17. Ivanova, N., Ahmadi, S., Chan, E., Fournier, L., Spagnolo, S., & Thompson, M. (2024). Detection of Ovarian Cancer Biomarker Lysophosphatidic Acid Using a Label-Free Electrochemical Biosensor. *Electrochem*, *5*(2), 243-257. https://doi.org/10.3390/electrochem5020015.
18. Bucur, C., Balescu, I., Petrea, S., Gaspar, B., Pop, L., Varlas, V., Stoian, M., Balalau, C., & Bacalbasa, N. (2024). Artificial Intelligence in Ovarian Cancers—From Diagnosis to Treatment; A Literature Review. *Journal of Mind and Medical Sciences*, *11*(2), 277-284. https://doi.org/10.22543/2392-7674.1531.
19. Du, Y., Wang, T., Qu, L., Li, H., Guo, Q., Wang, H., Liu, X., Wu, X., & Song, Z. (2024). Preoperative Molecular Subtype Classification Prediction of Ovarian Cancer Based on Multi-Parametric Magnetic Resonance Imaging Multi-Sequence Feature Fusion Network. *Bioengineering*, *11*(5), 472. https://doi.org/10.3390/bioengineering11050472.
20. Ziyambe, B., Yahya, A., Mushiri, T., Tariq, M. U., Abbas, Q., Babar, M., Albathan, M., Asim, M., Hussain, A., & Jabbar, S. (2023). A Deep Learning Framework for the Prediction and Diagnosis of Ovarian Cancer in Pre- and Post-Menopausal Women. *Diagnostics*, *13*(10), 1703. https://doi.org/10.3390/diagnostics13101703.
21. Jiang, X., Hu, Z., Wang, S., & Zhang, Y. (2023). Deep Learning for Medical Image-Based Cancer Diagnosis. *Cancers*, *15*(14), 3608. https://doi.org/10.3390/cancers15143608.
22. Ghoniem, R. M., Algarni, A. D., Refky, B., & Ewees, A. A. (2021). Multi-Modal Evolutionary Deep Learning Model for Ovarian Cancer Diagnosis. *Symmetry*, *13*(4), 643. https://doi.org/10.3390/sym13040643.
23. https://www.kaggle.com/datasets/bitsnpieces/ovarian-cancer-and-subtypes-dataset-histopathology